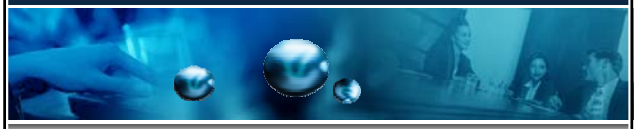
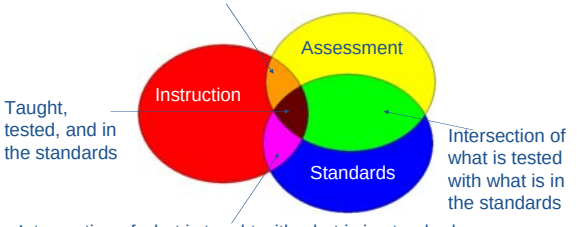
Practical Test Development Training



Assessment and Accountability Conference 2008

Alignment Relationships in Standards-Based Reform

Intersection of what is taught with what is tested.



Taught, tested, and in the standards

Intersection of what is tested with what is in the standards

Intersection of what is taught with what is in standards.

Smithson, J.L., Using Alignment to Analyze Standards, Assessments, & Classroom Practice, Mega-SCASS, Orlando, FL, Jan,

Validity

- All validity is of one kind, namely, "construct validity."
- "…constructs represent our best, albeit imperfect and fallible, efforts to capture the essence of traits that have a reality in behavior independent of our attempt to characterize them…"

Messick, S. (1998) Test validity: a matter of consequence. Social Indicators Research, 45, 35-44.

Construct Validity

- "…every test still underrepresents its construct to some degree and contains irrelevant variance, if for no other reason that it is a test and not a criterion performance…"

How To Know if a Test Is Aligned

- **Does the test's content match the content (topics and skills) in the standards?**
  - Each test item should correspond to an objective in the standards.
  - Key ideas in the standards should appear on the tests.

Standards and Tests: Keeping Them Aligned, Essential Information for Policy Makers, AERA, Vol. 1, Issue 1, Spring 2003,.

How To Know if a Test Is Aligned

- Do the tests and standards cover a comparable "range" or breadth of knowledge? Is there an appropriate "balance" of knowledge across the standards?
- Alignment studies look at whether a test fairly and effectively samples across the range of objectives described in the standards instead of focusing on only a few objectives or disproportionately sampling certain standards

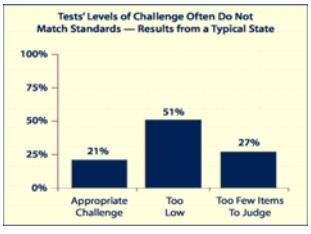
How To Know if a Test Is Aligned

- Does the level of cognitive demand or challenge called for in the standards match that required for student to do well on the assessment?
  - For example, if the standards require students to synthesize information and explain their thinking, but the test items only ask students to recall facts, the standards and test would not be well aligned.

How To Know if a Test Is Aligned

- Does the test avoid adding material that is irrelevant to the standard supposedly being assessed?
  - For example, a test item may have an inappropriate "source of challenge," requiring a student to read and understand a long passage about space travel, when it is seeking to measure a student's knowledge of how to estimate distances and travel times.

How well are tests aligned?



Standards and Tests: Keeping Them Aligned, Essential Information for Policy Makers
http://www.aera.net/uploadedFiles/Journals_and_Publications/Research_Points/RP_Spring03.pdf

What gets left off?

- Many assessments focus disproportionately on simpler standards.
  - A five-state review of English and math standards and tests by Achieve, Inc., concluded "The most challenging standards and objectives are the ones that are undersampled or omitted entirely . . . [and those] that call for high-level reasoning are often omitted in favor of much simpler cognitive processes."

Broad or Narrow Standards?

- "In some instances, state standards may not be sufficiently specific to allow an assessment to be aligned tightly with them or to provide adequate guidance for teachers."
- "(researchers)…have found that it is harder for tests to reflect the full range of knowledge included in state standards if content expectations are spread across a large number of standards."

Standards and Tests: Keeping Them Aligned, Essential Information for Policy Makers, AERA, Vol. 1, Issue 1, Spring 2003,.

Broad or Narrow Standards?

Standard H3 History of Michigan (Beyond Statehood)
Grade 4-H3.0.2 Use primary and secondary sources to explain how migration and immigration affected and continues to affect the growth of Michigan.
Grade 4-H3.0.3 Describe how the relationship between the location of natural resources and the location of industries (after 1837) affected and continues to affect the location and growth of Michigan cities.

Broad or Narrow Standards?

Standard H3 History of Michigan (Beyond Statehood)
Grade 4-H3.0.7 Use case studies or stories to describe the ideas and actions of individuals involved in the Underground Railroad in Michigan and the Great Lakes region.
Standard C1 Purposes of Government
Grade 4 - C1.0.3 Describe the purposes of government as identified in the Preamble of the Constitution.

Broad or Narrow Standards?

Students will…
R.CM.07.01 … connect personal knowledge, experiences, and understanding of the world to themes and perspectives in text through oral and written responses.

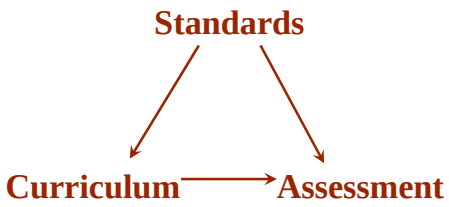
R.CM.07.02 … retell through concise summarization grade-level narrative and informational text.

Broad or Narrow Standards?

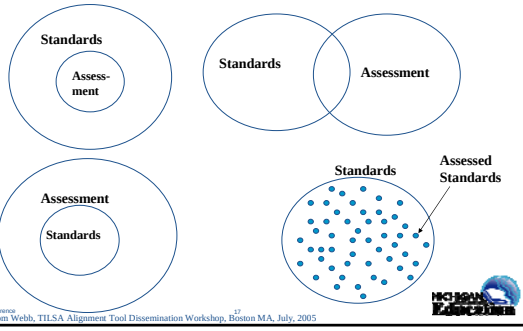
Students will…
R.CM.07.03 … analyze global themes, universal truths, and principles within and across texts to create a deeper understanding by drawing conclusions, making inferences, and synthesizing.

R.CM.07.04 … apply significant knowledge from grade-level science, social studies, and mathematics texts.

Alignment – A Simple Model

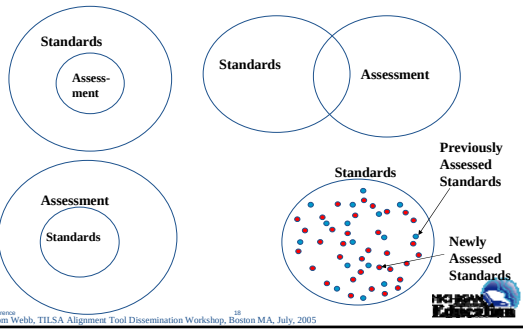


Alignment – A Simple Model



From Webb, TILSA Alignment Tool Dissemination Workshop, Boston MA, July, 2005

Alignment – A Simple Model



From Webb, TILSA Alignment Tool Dissemination Workshop, Boston MA, July, 2005

Alignment a la Webb

The degree to which expectations and assessments are in agreement and serve in conjunction with one another to guide the system toward students learning what is expected.

From Webb, Issues Related to Judging the Alignment of Curriculum Standards and Assessments, A

Alignment Criteria

- Categorical Concurrence
  - The assessment includes at least six items measuring content from each standard
- Depth-of-Knowledge Consistency
  - At least 50% of the items corresponding to a standard are at or above the level of knowledge of the standard

From Webb, Web Alignment Tool Training Draft Version 1.1, July, 2005

Alignment Criteria

- Range-of-Knowledge Correspondence
  - Fifty percent of the benchmarks for a standard had to have at least one related assessment item
- Balance of Representation
  - Items/activities are distributed among all of the benchmarks at least to some degree

From Webb, Web Alignment Tool Training Draft Version 1.1, July, 2005

Alignment Criteria

- Source of Challenge
  - This criterion is met if the primary difficulty of the assessment items is significantly related to students' knowledge and skill in the content area as represented in the standards

From Webb, Web Alignment Tool Training
Draft Version 1.1, July, 2005

Depth of Knowledge - Reading

- Level 1
  - requires students to receive or recite facts or to use simple skills or abilities
- Level 2
  - requires both comprehension and subsequent processing of text or portions of text

From Webb, Web Alignment Tool Training
Draft Version 1.1, July, 2005

Depth of Knowledge - Reading

- Level 3
  - Students may be encouraged to explain, generalize, or connect ideas
- Level 4
  - Higher-order thinking is central and knowledge is deep. The standard or assessment item at this level will probably be an extended activity with extended time provided for completing it

From Webb, Web Alignment Tool Training
Draft Version 1.1, July, 2005

Mathematics Level 1 DOK

- Identify
- Recall
- Recognize
- Use
- Measure

Mathematics Level 2 DOK

- Classify
- Organize
- Estimate
- Make observations
- Collect and display data
- Compare data

Mathematics Level 3 DOK

- Draw conclusions
- Cite evidence
- Develop a logical argument
- Explain phenomena in terms of concepts
- Apply concepts to solve a complex problem

Mathematics Level 4 DOK

- Requirements:
  - –Complex reasoning
  - –Planning
  - –Developing
  - –Thinking
- Usually implies work done over an extended period of time

Social Studies Level 1 DOK

- Recall
  - –Who
  - –What
  - –Where
  - –When
- Recite
- Reproduce information

Social Studies Level 2 DOK

- Contrast
- Compare
  - –People
  - –Places
  - –Events
  - –Concepts
- Describe or classify

Social Studies Level 3 DOK

- Draw conclusions
- Cite evidence
- Use concepts to explain how and why
- Solve problems
- Analyze similarities and differences

Social Studies Level 4 DOK

- Plan
- Investigate
- Develop
- Will most likely require an extended period of time

Developing Aligned Assessments

- Identify DOK of individual content standards
- Train item writers in alignment analysis
  - Identify which skills from content standards are used to correctly answer the item
  - Identify the DOK required to correctly answer the item

Developing Aligned Assessments

- Independently review items written for the assessments
  - Alignment to content standards
  - Alignment to DOK ratings for the content standards
- Construct Test to blueprint with items that have independently been aligned to the content standards and DOK ratings

Sample Test Blueprint

Number and Operations – Grade 4		
Domain	**GLCEs**	**Items**
Understand and use number notation	3	6
Use factors and multiples	4	6
Add and subtract whole numbers	1	6
Multiply and divide whole numbers	6	6

Grade 3: M.PS.03.03

- Solve applied problems involving money, length and time.
- Target DOK = 2
  - Applied
  - Not naked math

M.PS.03.03 Item – DOK = 1

John put 4 quarters in his bank. How much money did he put in his bank?

A. $1.00
B. $1.25
C. $2.00
D. $4.00

Easily solved by recall.

M.PS.03.01 Item – DOK = 2

Mr. Garza has two boards. One board is 80 cm long and the other is 40 cm long. What is the total length of the two boards?

A. 1 meter 20 centimeters
B. 1 meter 40 centimeters
C. 2 meters
D. 2 meters 20 centimeters

Organize data & use arithmetic operations

M.PS.03.01 Item – DOK = 3

Which of these babies is *oldest*?

A. Mary is 13 months old.
B. Arthur is 1 year 2 months old.
C. Jeanne is 9 months old.
D. Patty is 1 years old.

Apply concepts to solve a complex problem

Grade 7: A.FO.07.03

- From applied situations, generate and solve linear equations of the form ax + b = c and ax + b = cx + d, and interpret solutions.
- Target DOK = 2
  - –Apply

A.FO.07.03 Item – DOK = 1

Which of the following is another way to represent this expression?

$3x + (2x - 1) + 5(x + 2)$

A. $10x + 4$

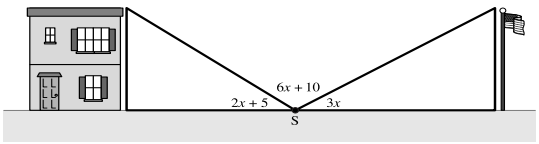
B. $10x + 6$

C. $10x + 9$

D. $11x + 4$

A.FO.07.03 Item – DOK = 2

A surveyor stood at point S and measured the angles indicated below.



A.FO.07.03 Item – DOK = 2

Which of the following equations could be used to find the number of degrees in each of these three angles?

Number of degrees in a straight line = 180 degrees.

A. $11x + 15 = 180$

B. $11x + 50 = 180$

C. $12x + 15 = 180$

D. $36x + 50 = 180$

A.FO.07.03 Item – DOK = 3

An elevator can hold a maximum of 10 people who weigh an average of 180 pounds each. If a box of freight weighing 275 pounds is placed on the elevator, what strategy can be used to determine the number of people of average weight who can safely get on the elevator?

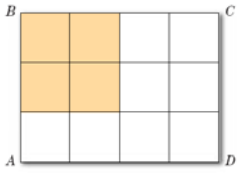
A. Solve $275 + 180x = 1800$ and round down to the next whole number.

B. Solve $180x = 1800 + 275$ and round up to the next whole number.

C. Solve $275 + 180x = 1800$ and do not round.

D. Solve $1800 + 180x = 275$ and do not round.

Small changes in test items …



What fraction of ABCD is shaded?

Requires less thinking because it explicitly defines the square units to make the shaded area easier to measure.

… can shift cognitive demand

What fraction of ABCD is shaded?

Increases the challenge by requiring that the student make measurements to determine what portion of the figure is shaded.

Test Development Process

- Step 1 – Define what you want to assess
  - –Blueprint
  - –Items per concept/unit
  - –Overall length of the test

Test Development Process

- Step 2 – Try out the test
  - –Colleague
  - –Other students
  - –Are the results as you expected?

Test Development Process

- Step 3 – Use the test with its intended population
  - –Compute score distribution
    - Is it a normal distribution?
  - –Set standards
    - What is “Passing?” “Exemplary?”

Test Development Process

- Step 4 - Make notes of your observations
  - –Unexpected responses
  - –Did your better students score higher?
  - –What items can you keep for future use?

Contacts

- Bill Brown
  brownb6@michigan.gov
  (517) 335-0568
- Tom Wessels
  twessels@tbaisd.k12.mi.us